# Earned Schedule 10 Years Later

## Analyzing Military Programs

**Kevin T. Crumrine, Air Force Operations C2 Division**
**Jonathan D. Ritschel, Ph.D., Air Force Institute of Technology**
**Edward White, Ph.D., Air Force Institute of Technology**

**Abstract.** It has been 10 years since Walt Lipke first introduced the concept of Earned Schedule (ES). While progress has been made in understanding the utility of ES in some small scale and limited studies, a significant analysis of ES in DoD acquisition programs is missing. This paper first analyzes whether ES and Earned Value Management (EVM) provide fundamentally different information for program managers. It then examines which technique, ES or EVM, provides more timely and accurate schedule predictors in a broad spectrum of military weapon system programs. We find ES to be more timely and accurate both in software intensive contracts and in the sample size as a whole.

### Background

EVM has been the premier method of program management and program cost forecasting within the DoD since its inception in the 1960s. However, there are well-documented limitations to EVM particularly with respect to schedule analysis [1]. These limitations include: 1) reporting schedule variance in terms of dollars rather than time 2) the regression of EVM schedule efficiency metrics (SPI($)) to 1 as projects near completion, despite variable schedule performance and 3) the regression of EVM schedule variance metrics (SV($)) to zero as projects near completion. For practitioners in the field, these issues make traditional EVM schedule analysis unwieldy. To mitigate these limitations, Walt Lipke developed the concept of ES as an alternative to EVM [1]. Lipke's ES construct measures schedule performance with analogous earned value metrics dubbed Schedule Performance Index (SPI(t)) and Schedule Variance (SV(t)) where (t) indicates the metric is reported in time.

But the question remains: Should DoD managers utilize ES as a preferred schedule analysis technique? Program managers should only implement ES analysis as part of their tool kit if it provides additional benefit beyond the established EVM techniques. Thus, the answer to the question becomes an empirical matter. Previous studies (Henderson [2] [3], Lipke [4], Vanhoucke & Vandevoorde [5], Rujirayanyong [6], Tzaveas, Katsavounis & Kalfakakou [7], Lipke [8]), have examined the efficacy of ES, but these studies were all limited by their extremely small sample size or lack of relevance to the DoD.

This paper overcomes the previous literature shortcomings by analyzing over 64 contracts in major Air Force aircraft acquisition programs to determine whether ES provides more timely and accurate information. These contracts include software intensive contracts such as avionics along with hardware intensive contracts such as engines, capturing the full spectrum of an aircraft acquisition effort. The large sample size and direct relationship to military programs makes the results of this analysis directly applicable to DoD software and hardware program managers.

### Data Source

The data for this analysis is from the Defense Acquisition Management Information Retrieval (DAMIR) system. DAMIR is comprised of all Contractor Performance Report (CPR) data for major DoD acquisition programs. The CPR data contains the monthly and quarterly performance information derived from the contractors EVMS system for all Work Breakdown Structures (WBS) within each contract of a program. Thus, it provides the cost and schedule status for the contract [9].

This analysis focuses on 64 Acquisition Category (ACAT) 1 aircraft contracts at the summary level (WBS 1). The programs comprising the dataset have completed their acquisition phase, and are either in their operational phase, or have been retired from the Air Force fleet. The 64 contracts result in 1,087 data points in the full analysis. We specifically examine the software intensive avionics contracts as a group, in addition to an aggregated analysis of all 64 contracts.

### Methodology and Results

### Preliminary Analysis

The first question to answer is whether ES and EVM provide fundamentally different information to program managers. Once this is ascertained, the method that provides better information, measured in this paper by timeliness and accuracy, can be determined. We statistically test the difference between ES and EVM through a paired t-test of SPI($) and SPI(t). A paired t-test measures the mean difference between two sets of numbers. The null hypothesis is that there is no difference between the methods. Table 1 shows the results.

| t-Test: Paired Two Sample for Means | | |
|---|---|---|
| | | |
| | *Variable 1* | *Variable 2* |
| Mean | 0.939165476 | 0.95750293 |
| Variance | 0.008831643 | 0.006653895 |
| Observations | 1087 | 1087 |
| Pearson Correlation | 0.689419981 | |
| Hypothesized Mean Difference | 0 | |
| df | 1086 | |
| t Stat | -8.623145392 | |
| P(T<=t) one-tail | 1.13734E-17 | |
| t Critical one-tail | 1.646257934 | |
| P(T<=t) two-tail | 2.27467E-17 | |
| t Critical two-tail | 1.962150792 | |

*Table 1: Paired t-test SPI($) vs SPI(t)*

As shown in Table 1, the p-value of the t-test is 2.27E-17, well below our significance level of 0.05. Therefore, the null hypothesis is rejected. This means there is a statistically significant likelihood that ES and EVM information are fundamentally different from each other. In practical terms, this indicates that utilizing the ES technique provides additional information to the program manager. The question then becomes whether the ES information is more valuable, as measured by its timeliness and accuracy.

# Report Documentation Page

*Form Approved*
*OMB No. 0704-0188*



1. REPORT DATE
**APR 2014**

2. REPORT TYPE

3. DATES COVERED
**00-00-2014 to 00-00-2014**

4. TITLE AND SUBTITLE
**Earned Schedule 10 Years Later Analyzing Military Programs**

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
**Air Force Institute of Technology,2950 Hobson Way ,Wright Patterson AFB,OH,45433**

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
**Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT
**It has been 10 years since Walt Lipke first introduced the concept of Earned Schedule (ES). While progress has been made in understanding the utility of ES in some small scale and limited studies, a significant analysis of ES in DoD acquisition programs is missing. This paper first analyzes whether ES and Earned Value Management (EVM) provide fundamentally different information for program managers. It then examines which technique, ES or EVM, provides more timely and accurate schedule predictors in a broad spectrum of military weapon system programs. We find ES to be more timely and accurate both in software intensive contracts and in the sample size as a whole.**

15. SUBJECT TERMS

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | |
| **unclassified** | **unclassified** | **unclassified** | **Same as Report (SAR)** | **4** | |

## Testing Timeliness

Metrics help managers determine when a problem is occurring so that corrective action may be taken. For this analysis, a problem was defined as a SPI($) or SPI(t) < 0.90. The intent of this test is to determine whether EVM or ES is an earlier detector of problems in meeting program schedule objectives.[1]

The initial dataset examined is the subset of software intensive avionics contracts. Of these contracts that both ES and EVM identify as a problem, EVM identifies the problem at the 18.87% completion point, while ES identifies the problem at the 16.88% completion point. EVM, therefore, detects about 2% earlier than ES. However, drawing conclusions based on this is misleading. Rather the analysis necessitates that we look at all the avionics contract problems detected, even if only one of ES or EVM detects it. See Figure 1.

Figure 1 shows that ES strictly dominates EVM. ES identifies more problems at every completion point of the contract. More importantly, at the earlier stages of the program, ES detects more problems. For instance, at the 20% completion point, ES detects seven programs with problems while EVM only detects two. This early difference in detection is critical as it allows program mangers to take corrective action early in the program. Figure 1 also demonstrates a second area where ES is more valuable than EVM. Note that around the 2/3 program completion point, EVM no longer detects any problems, while ES remains useful in problem detection through the end of program completion.

Next we analyzed the full 64-contract dataset. The total number of SPI(t) and SPI($) values below 0.90 were analyzed at each of the following program completion points: 20%, 40%, 50%, 60%, 80%, and 90%. See Table 2.

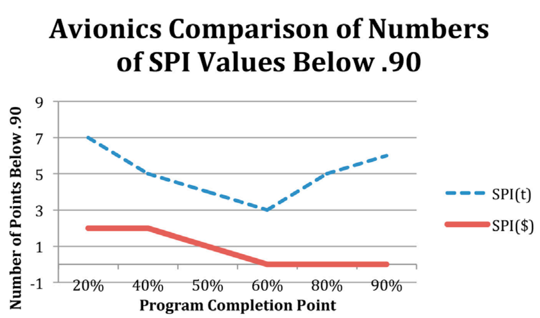



*Figure 1: Avionics Comparison of Numbers of SPI Values Below 0.90*

| | 20% | 40% | 50% | 60% | 80% | 90% |
|---|---|---|---|---|---|---|
| SPI(t) | 20 | 17 | 11 | 14 | 15 | 20 |
| SPI($) | 12 | 11 | 4 | 5 | 2 | 1 |

*Table 2: Number of SPI Value Below 0.90 Over Time*

Table 2 shows quite clearly that as early as the 20% program completion point, the ES metric was indicating a problem more frequently than the EVM metric. Additionally, this gulf in detection exacerbates over the life of the program, consistent with previous literature: as a contract approaches its completion point, EVM yields an SPI($) value that approaches 1.0, indicating that the program is on schedule even if it is not. This is seen at the 90% completion point where SPI(t) correctly found 20 programs to be "in trouble," while SPI($) found only 1.

### Testing Accuracy

Two analyses are performed to compare the accuracy of ES and EVM. First, we measure the SPI($) and SPI(t) in relation to the final schedule result. Whichever method is closer to the final contract over/under run is deemed to be the more accurate technique. The results for the avionics subset of contracts are shown in Table 3.

| | Number of Occurrences | Percentage of Overall Occurrences (%) |
|---|---|---|
| Earned Value Management | 107 | 43.67 |
| Earned Schedule | 126 | 51.43 |
| EVM = ES | 12 | 4.90 |

*Table 3: Accuracy of ES and EVM in Avionics Contracts*

Table 3 shows that ES is more accurate than EVM in the avionic subset. There is approximately an 8% difference between the techniques for these software intensive contracts. While this finding is significant, the accuracy margin widens to 21% when the full 64-contract dataset is analyzed. Of the 1,087 data points, EVM is closer to the final schedule result 37% of the time, while ES is the more accurate technique 58% of the time. The EVM and ES values are equivalent 5% of the time. Thus, for both the avionics subset and the dataset as a whole, ES trumps EVM in accuracy.

The second analysis, shown in Figure 2, depicts the frequency of contracts having a particular percentage of their data points closer to the final schedule result. For instance, the B1B Offensive Avionics Lot 1 has 15 points where the SPI(t) is closer to the final schedule result than the SPI($). There are 20 data points for this program, so ES is closer to the final schedule result 75% of the time. As depicted in Figure 2, this contract is 1 of 9 contracts where the SPI(t) value is closest to the final schedule result between 70% and 75% of the time. There is a definite skew left to this histogram, demonstrating the greater accuracy of ES. In fact, there are only four programs that have less than 30% of their data points with SPI(t) values closer to the final schedule result.

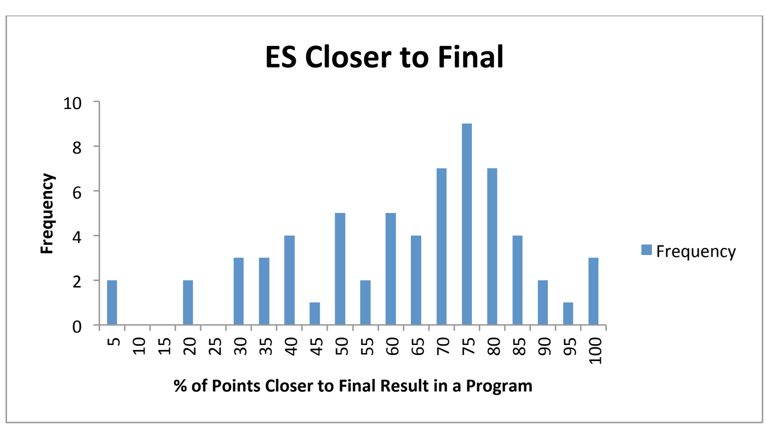



*Figure 2: Distribution of Programs With ES Closer to Final Program Delivery*

In addition to analyzing the contracts at an individual level, we also want to determine how the entire portfolio acts over a period of time. As shown in Figure 3, the ES metric dominates the EVM metric at all program completion percentage points. This result points to ES providing valuable information to the program manager.

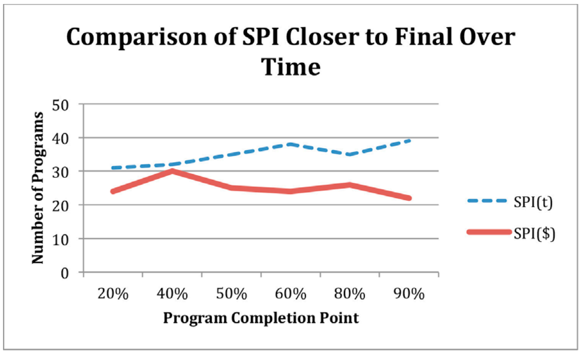



*Figure 3: Comparison of SPI Closer to Final Over Time*

### Other Schedule Techniques: the Critical Path

EVM is not the only technique used by DoD program mangers to analyze schedule. The most common methodology is the Critical Path Method. Lipke [4] argues that Earned Schedule is applicable to the critical path. We examine this finding in a small subset of our data. Our results show a fundamental disconnect between the level of Earned Value data collected and the level of Critical Path data utilized by the program offices. Specifically, we find that earned value data is collected at a much higher level than the level in which critical path analysis is being performed, rendering a comparison infeasible. This does not necessarily suggest that ES is inapplicable to the CPM in the DoD. Rather, it points to the necessity of making contractor EVMS reporting at a lower level as part of contract deliverables than is typical today. This, of course, would result in increased contract costs. More research is needed in this area to determine that cost/benefit ratio.

### Conclusion

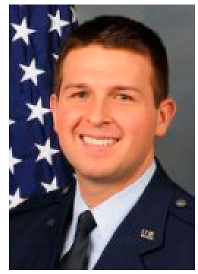
This paper has demonstrated with statistical significance that ES is fundamentally different from EVM. Our empirical analyses of 64 contracts show that not only is there a difference between the two techniques, but that difference is wide enough to warrant a reconsideration of the use of ES in DoD programs. Specifically, we find ES to be both timelier and more accurate than traditional EVM schedule analysis.

The practical implications of our research are straightforward. Due to our inability to thoroughly test ES against CPM, we stop short of recommending ES as its replacement. However, our analysis indicates ES warrants more intensive use for schedule analysis in DoD programs. Specifically, based on the findings of our research, we believe that DoD ACAT I programs should embrace ES as a complementary tool (i.e. the primary cross-check) to the CPM method that is predominately utilized. Traditional EVM schedule analysis techniques should not be abandoned completely, but should be secondary to the CPM and ES techniques.

### Disclaimers:

The views expressed in this article are those of the authors and do not necessarily reflect the views of the DoD or any of its agencies.

## ABOUT THE AUTHORS

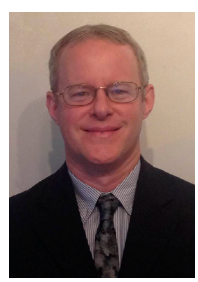



Kevin T. Crumrine is a Captain in the United States Air Force. He graduated from the U.S. Air Force Academy in 2008 with a B.S. in Management, and from the Air Force Institute of Technology in 2013 with a M.S. in Cost Analysis. Kevin currently serves the Deputy Cost Chief for the Operations C2 Division at Hanscom AFB, MA.

**E-mail: kevin.crumrine@gmail.com**

Jonathan D. Ritschel, Ph.D., is an Assistant Professor and Director Cost Analysis Program in the Department of Systems Engineering and Management at the Air Force Institute of Technology (AFIT). He received his BBA in Accountancy from the University of Notre Dame, his M.S. in Cost Analysis from AFIT, and his Ph.D. in Economics from George Mason University. Lt Col Ritschel's research interests include public choice, acquisition reform effects on cost growth, and economic institutional analysis.

**E-mail: jonathan.ritschel@afit.edu**



Edward "Tony" White, Ph.D. is an associate professor of Statistics at the Air Force Institute of Technology. His research interests include design of experiments, biostatistics, growth curves, linear and nonlinear regression, categorical data analysis, log-linear models, statistical simulation, and response surface modeling. White received his Ph.D. in statistics from Texas A&M University in 1998.

**Phone: 937-255-3636 x4540**
**E-mail: edward.white@afit.edu**

## REFERENCES


1. Lipke, Walt. "Schedule is Different," The Measurable News 3 (2003): 10-15.
2. Henderson, Kym. "Earned Schedule: A Breakthrough Extension to Earned Value Theory? A Retrospective Analysis of Real Project Data," The Measurable News (2003): 13-23.
3. Henderson, Kym. "Earned Schedule in Action," The Measurable News 8 (2005): 23-30.
4. Lipke, Walt. "Applying Earned Schedule to the Critical Path and More," The Measurable News (2006): 26-30.
5. Vanhoucke, M., S. Vandevoorde. "Measuring the Accuracy of Earned Value / Earned Schedule Forecasting Predictors," The Measurable News, Winter (2007a): 26-30.
6. Rujirayanyong, Thammasak. "A Comparison of Three Completion Date Predicting Methods for Construction Projects," Journal of Research in Engineering and Technology, Volume 6, (2009).
7. Tzaveas, Theodoros, Stefanos Katsavounis, and Glikeria Kalfakakou. "Analysis of Project Performance of a Real Case Study and Assessment of Earned Value and Earned Schedule Techniques for the Prediction of Project Completion Date." Proceedings of IPMA Conference, May 2010 (Crete, Greece).
8. Lipke, Walt. "Earned Schedule Application to Small Projects," The Measurable News 2 (2011): 25-31.
9. Defense Contract Management Agency. Department of Defense Earned Value Management Implementation Guide. October 2006. Alexandria, VA: DCMA, 2006. 10 (2012) <http://guidebook.dcma.mil/79/EVMIG.doc>


## NOTES

1. Preliminary data analysis demonstrated that there are frequent occurrences where a program's SPI value drops below 0.90 early in a program and quickly recovers. This led to the potential for false conclusions, necessitating a different analysis. Therefore, to be counted as "detecting" a problem in our analysis, the SPI metric must remain below 0.90 for multiple consecutive time periods.